مجلة جامعة فزان العلمية
Fezzan University scientific Journal
journal@fezzanu.edu.ly

Deep Learning-Based Bone Tumor Detection from X-ray Images: A DenseNet121 Study

*Abdelkader Alrabai[1]
[1]Physics Department – Faculty of Education – Wadi Alshatti University

Abstract
The identification of bone tumors from radiographic images is an important task in medical diagnosis, as early detection can support timely clinical decision-making and improve patient management. This study examines the use of the DenseNet121 deep learning architecture for automated bone tumor detection using X-ray images. The task is formulated as a binary classification problem, in which radiographs are categorized as either normal or tumor cases, with the aim of exploring the feasibility of deep learning as a supportive diagnostic tool.
The DenseNet121 model was evaluated on the BTXRD dataset, which contains labeled bone X-ray images representing both healthy and tumor-affected cases. Model performance was assessed using standard evaluation metrics to provide a balanced and objective analysis of its classification behavior. These metrics offer insight into both overall accuracy and the model's discriminative ability across different thresholds.
The experimental results indicate moderate performance. The model achieved an accuracy of 75.73%, reflecting a reasonable ability to differentiate between normal and tumor images, although some misclassifications remain. The area under the curve (AUC) reached 83.87%, suggesting that the model maintains good class separability despite the moderate accuracy.
Overall, these findings demonstrate the potential of DenseNet121 for bone tumor detection from radiographs, while also highlighting the need for further improvements through larger datasets, enhanced feature learning, and more comprehensive validation before clinical application.
Keywords: Bone tumors, DenseNet121, Radiographic.

Introduction

Bone tumors include a wide range of growths that can start in the bone itself or spread there from other parts of the body. Those that begin in the bone are called primary tumors, while ones that spread from elsewhere are known as secondary or metastatic tumors. Primary bone tumors can be either benign or malignant, and each type tends to behave differently, show distinct symptoms, and respond to treatment in its own way [1].
Imaging is essential in evaluating suspected bone tumors—from spotting the lesion to judging its potential seriousness and spread. X-rays are usually the first step and can reveal important clues about the tumor's nature, especially when combined with the patient's history. More advanced scans, such as CT, MRI, and nuclear imaging, help define how far the tumor has spread and guide further treatment. Key details, such as the edges of the lesion, bone changes, tissue involvement, and overall

*Corresponding author: a.alrabai@wau.edu.ly Received in 12 January 2026
Revised form 26 February 2026 Accepted 20 march 2026 - Publication date 26 march 2026

appearance, help estimate how aggressive the tumor might be and shape the next steps in care [2].

Identifying bone tumors early is crucial for preserving limb function and lowering the risk of serious complications or death. Since these tumors are uncommon, diagnosis typically relies on a mix of symptoms, imaging tests, and pathology. However, this process can be slow, costly, and heavily influenced by the clinician's level of experience, making timely and precise diagnosis challenging. This highlights the pressing need for faster, more dependable diagnostic approaches in everyday clinical settings [3].

Bone and soft tissue tumors make up only a small fraction of all cancers. Because they are uncommon, accurate interpretation of imaging often depends on radiologists who specialize in musculoskeletal tumors. To improve diagnostic accuracy and reduce the risk of missing serious cases, efforts over recent years have explored using AI tools in clinical practice.[4].

Deep learning, a type of machine learning built on neural networks, allows computers to automatically learn patterns from data without relying on manually crafted features. This approach supports end-to-end model development and is especially effective for handling large, complex datasets. In medical imaging, deep learning can assist in identifying abnormalities, supporting diagnoses, and minimizing errors. It has already shown strong results in various medical applications [5]. Deep learning has been applied in musculoskeletal imaging for tasks such as identifying different tissues, improving image quality, and spotting signs of disease. When used for detecting conditions, it can be particularly helpful for radiologists by boosting accuracy and minimizing mistakes that might occur due to fatigue or distraction [6].

Recent progress in deep learning has led to notable improvements in medical image analysis, especially in the automated detection and classification of bone tumors from radiographic images. A growing body of research has addressed this topic, with several studies [7–12] examining the effectiveness of different deep learning architectures for identifying and categorizing bone tumors in X-ray data.

Building upon previous research, this study investigates the automatic classification of radiographic bone images into two categories: normal and tumor. The proposed approach utilizes the DenseNet121 deep learning architecture to exploit its dense connectivity mechanism, which facilitates efficient feature propagation and improves the learning of discriminative patterns within X-ray images. By leveraging these capabilities, the model aims to enhance classification performance and contribute to more reliable computer-aided diagnostic systems. The study specifically focuses on the binary classification of bone X-ray images to identify the presence or absence of bone tumors. This simplified classification framework represents an essential preliminary stage toward the development of more advanced automated systems for bone tumor detection and analysis. Early identification of abnormal bone structures through such models may assist clinicians in screening large volumes of radiographic data and improving diagnostic efficiency. Therefore, the proposed method may provide a supportive tool for medical professionals by facilitating early detection and aiding in clinical decision-making.

Method

The methodology of this study follows a structured pipeline, as illustrated in Figure 1. The DenseNet121 [13] architecture is adopted to identify bone tumors in radiographic images by assigning each image to one of two classes: healthy bone or tumor-affected bone.

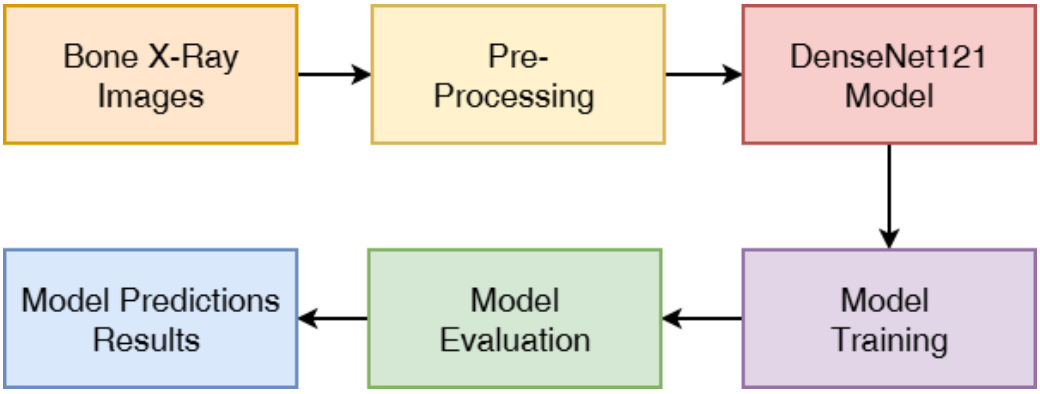

Figure 1: Workflow of the method

The experiments are conducted using the BTXRD [14] dataset, a publicly accessible repository of bone tumor X-ray images comprising a total of 3,746 radiographs. Among these, 1,879 images represent normal bone structures, while 1,867 images correspond to tumor cases, including both benign and malignant conditions.
The BTXRD dataset consists of X-ray images of different anatomical regions of the body, including the upper limbs, lower limbs, and pelvic area, with both normal and tumor cases represented. These images were captured from multiple radiographic views, such as frontal, lateral, and oblique perspectives. Figure 2 presents representative examples from both the healthy bone and tumor-affected categories.

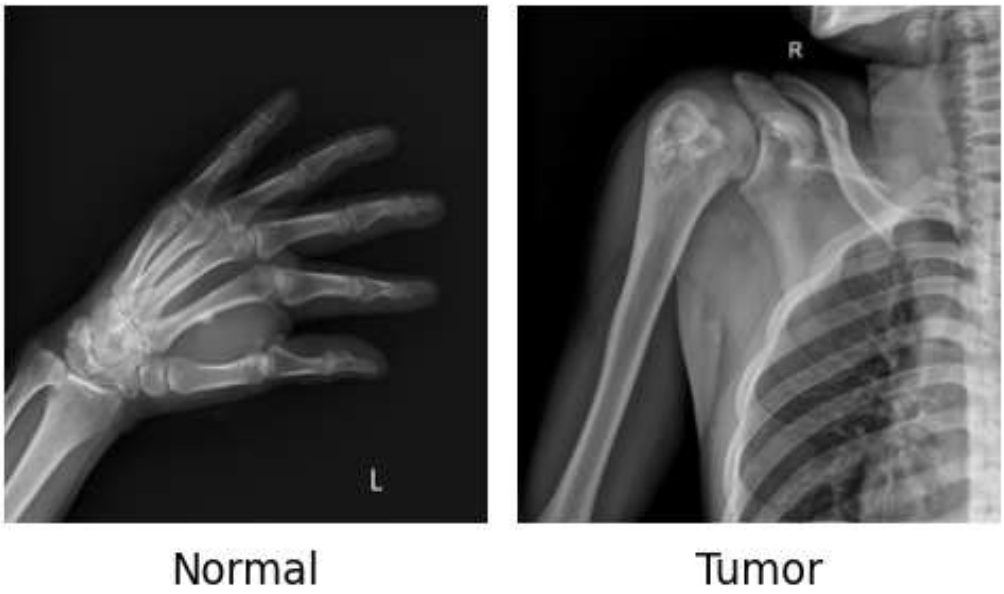

Figure 2: Tumor and normal image samples

Prior to model training, a preprocessing step was applied in which all radiographic images were resized to 224 × 224 pixels and normalized to achieve a uniform intensity distribution across the dataset. To ensure a reliable assessment of the employed model, the dataset was partitioned into training, validation, and testing sets using a ratio of 70%, 20%, and 10%, respectively.
The DenseNet121 model was configured for binary classification and trained using the Adam optimizer with a learning rate of 0.0001 and a batch size of 32 over 25

training epochs. The performance of the developed model was evaluated using standard performance metrics to quantitatively assess its ability to correctly classify the input images.

Results and Discussion

The model was trained using X-ray images belonging to normal and tumor classes, and its performance was subsequently assessed on an independent test set to evaluate its generalization capability. Figure 3 summarizes the model's training behavior, showing steady convergence with improving accuracy and decreasing loss over successive epochs. Training and validation accuracies increase rapidly at early stages and remain closely aligned, reaching about 75% and 74%, respectively, which indicates good generalization. The loss curves exhibit a consistent downward trend without signs of overfitting, as validation loss remains comparable to or slightly lower than training loss. Minor fluctuations in validation metrics suggest limited instability that could be reduced with learning rate scheduling. Overall, the model demonstrates stable and effective learning, with potential for slight further improvement through extended training.

In addition, Table 1 summarizes the key evaluation metrics of the DenseNet121 model on the test set for both normal and tumor classes. These metrics provide a comprehensive view of how well the model distinguishes between healthy and tumor-affected bone images.

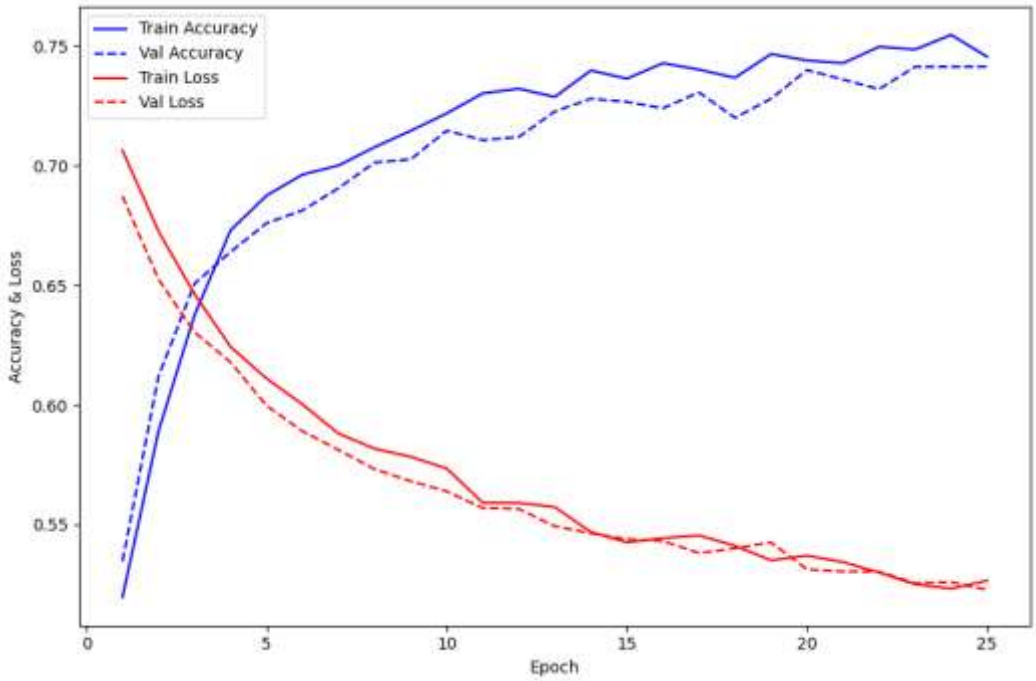

Figure 3: Training curves

Table 1. Performance metrics of DenseNet121 on test set

Metric\ Class	Precision	Recall	F1-score	Accuracy
Normal	0.7487	0.7766	0.7624	0.7573
Tumor	0.7667	0.7380	0.7520	

From Table 1, it is evident that the model maintains a balanced performance across both classes. For the normal class, the precision is 0.7487 and the recall is 0.7766, resulting in an F1-score of 0.7624. This indicates that the model is slightly better at

correctly identifying normal images than misclassifying them. For the tumor class, the precision is marginally higher at 0.7667, while the recall is slightly lower at 0.7380, yielding an F1-score of 0.7520. The overall accuracy of the model is 0.7573, reflecting its consistent ability to correctly classify the majority of X-ray images.
The close alignment of precision and recall values for both classes suggests that the model does not heavily favor one class over the other, which is particularly important in medical diagnosis to avoid false negatives or false positives. Overall, Table 1 demonstrates that the DenseNet121 architecture provides a reliable balance between sensitivity and specificity for detecting bone tumors in radiographs.
Furthermore, Figure 4 presents the confusion matrix for the DenseNet121 model's predictions on the test set, providing a visual representation of how well the model distinguishes between normal and tumor bone images. This figure complements the quantitative metrics in Table 1 by showing the actual counts of correct and incorrect classifications for each class.
From Figure 4, it is clear that the model correctly identified 146 normal images and 138 tumor images, while 42 normal images were misclassified as tumors and 49 tumor images were incorrectly labeled as normal. These results indicate that the model maintains a reasonably balanced performance across both classes, with slightly higher accuracy for normal bone images. The number of misclassifications is relatively low, suggesting that the model is effective at capturing the distinguishing features between healthy and tumor-affected bones. Figure 4 reinforces the reliability of the DenseNet121 model in the binary classification of bone X-ray images, demonstrating its potential utility in aiding clinical diagnosis.

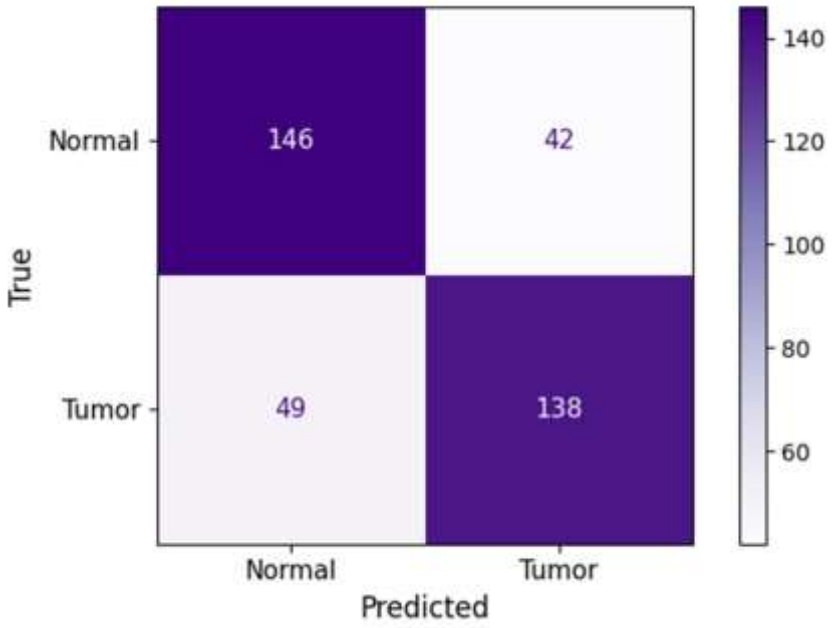

Figure 4: confusion matrix

Moreover, Figure 5 illustrates the ROC curve of the DenseNet121 model on the test dataset, offering insight into its discriminative capability across different classification thresholds. Unlike accuracy-based metrics, the ROC curve highlights the trade-off between the true positive rate and the false positive rate, making it particularly informative for evaluating medical classification models.

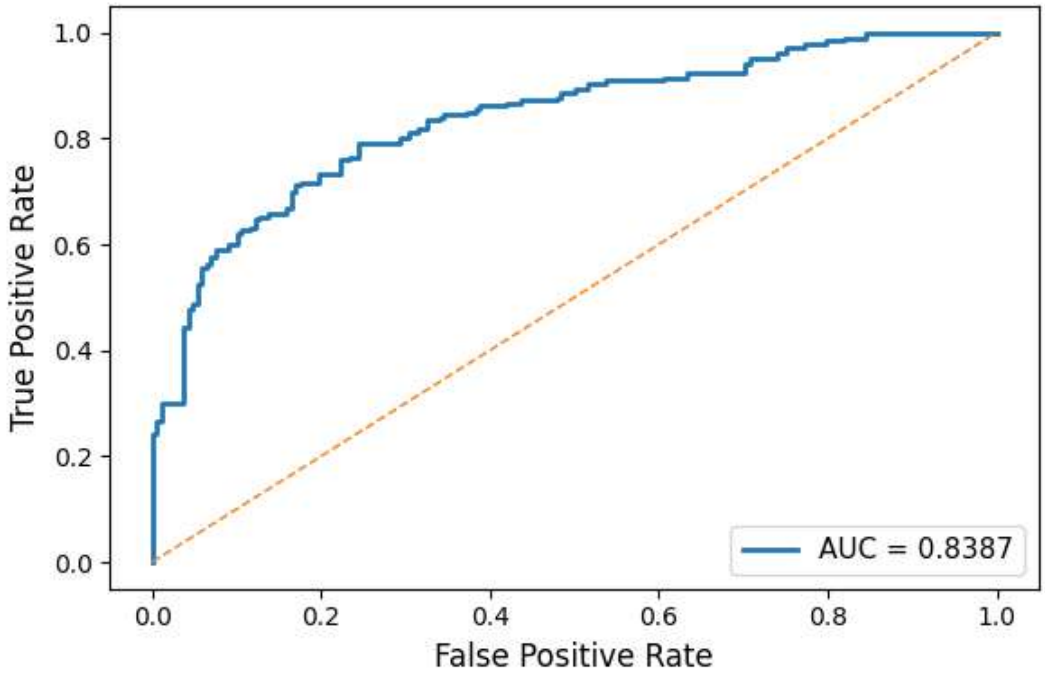

Figure 5: ROC curve

As shown in Figure 5, the ROC curve consistently remains well above the diagonal reference line. This diagonal represents random classification. The curve's behavior indicates that the model effectively separates tumor cases from normal bone images, even across a wide range of decision thresholds. The calculated AUC of 0.8387 reflects strong overall classification performance. This suggests the model has a high probability of correctly ranking tumor images higher than normal ones.

The relatively steep rise of the curve at lower false positive rates shows that the model achieves good sensitivity without a substantial increase in false alarms. This characteristic is especially important in clinical settings. In such contexts, early and accurate detection of tumors is critical. Overall, Figure 5 confirms that the DenseNet121 model exhibits robust discriminative power. This supports the reliability of results reported earlier in Table 1 and Figure 4.

Limitations

X-ray imaging remains a common choice for bone assessment due to its availability and low cost; however, it poses notable challenges for tumor detection. Limited soft-tissue contrast, overlapping structures, and non-specific radiographic features often reduce diagnostic precision. These constraints are further reflected in AI-based systems, where restricted and less diverse annotated datasets can limit robustness and generalization.

In this study, the DenseNet121 model achieved around 75% accuracy in classifying normal and tumor bone X-ray images, indicating moderate performance. While the model demonstrates an ability to learn meaningful patterns, the observed misclassifications suggest that additional optimization is required before clinical deployment.

The relatively small dataset of 3,746 images and the use of binary classification also limit the scope of the study. Rare tumor variations and clinically relevant distinctions, such as benign versus malignant cases, are not fully captured. Moreover, reliance on a single dataset and two-dimensional X-ray images restricts generalizability and may obscure critical tumor details. Future research should focus

on larger and more diverse datasets, finer-grained classification, the inclusion of clinical information, and explainable AI methods to enhance performance and clinical relevance.

References:

[1] Hosseini, H., Heydari, S., Hushmandi, K., Daneshi, S., & Raesi, R. (2025). Bone tumors: a systematic review of prevalence, risk determinants, and survival patterns. BMC cancer, 25(1), 321.

[2] Debs, P., Ahlawat, S., & Fayad, L. M. (2024). Bone tumors: state-of-the-art imaging. Skele-tal radiology, 53(9), 1783-1798.

[3] Zhou, X., Wang, H., Feng, C., Xu, R., He, Y., Li, L., & Tu, C. (2022). Emerging applica-tions of deep learning in bone tumors: current advances and challenges. Frontiers in oncolo-gy, 12, 908873.

[4] Vogrin, M., Trojner, T., & Kelc, A. P. R. (2020). Artificial intelligence in musculoskeletal oncological radiology. Radiology and oncology, 55(1), 1.

[5] Jiang, X., Hu, Z., Wang, S., & Zhang, Y. (2023). Deep learning for medical image-based cancer diagnosis. Cancers, 15(14), 3608.

[6] Kijowski, R., Liu, F., Caliva, F., & Pedoia, V. (2020). Deep learning for lesion detection, progression, and prediction of musculoskeletal disease. Journal of magnetic resonance imag-ing, 52(6), 1607-1619.

[7] Hinterwimmer, F., Serena, R. S., Wilhelm, N., Breden, S., Consalvo, S., Seidl, F., ... & Rueckert, D. (2024). Recommender-based bone tumour classification with radiographs—a link to the past. European Radiology, 34(10), 6629-6638.

[8] Breden, S., Hinterwimmer, F., Consalvo, S., Neumann, J., Knebel, C., von Eisenhart-Rothe, R., ... & Lenze, U. (2023). Deep learning-based detection of bone tumors around the knee in X-rays of children. Journal of Clinical Medicine, 12(18), 5960.

[9] Park, C. W., Oh, S. J., Kim, K. S., Jang, M. C., Kim, I. S., Lee, Y. K., ... & Seo, S. W. (2022). Artificial intelligence-based classification of bone tumors in the proximal femur on plain radiographs: System development and validation. PLoS One, 17(2), e0264140.

[10] Pan, C., Lian, L., Chen, J., & Huang, R. (2023). FemurTumorNet: Bone tumor classi-fication in the proximal femur using DenseNet model based on radiographs. Journal of Bone Oncology, 42, 100504.

[11] He, Y., Pan, I., Bao, B., Halsey, K., Chang, M., Liu, H., ... & Bai, H. X. (2020). Deep learning-based classification of primary bone tumors on radiographs: a preliminary study. EBioMedicine, 62.

[12] Xu, D., Li, B., Liu, W., Wei, D., Long, X., Huang, T., ... & Gao, Z. (2024). Deep learning-based detection of primary bone tumors around the knee joint on radiographs: a multicenter study. Quantitative Imaging in Medicine and Surgery, 14(8), 5420.

[13] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[14] Yao, S., Huang, Y., Wang, X., Zhang, Y., Paixao, I. C., Wang, Z., ... & Song, J. (2025). A Radiograph Dataset for the Classification, Localization, and Segmentation of Primary Bone Tumors. Scientific Data, 12(1), 88.